Evolution of a Performance Metric for Urban Search and Rescue Robots (2003)

Adam Jacoff, Brian Weiss, and Elena Messina

Intelligent Systems Division
National Institute of Standards and Technology

ABSTRACT

This paper discusses the advancement and proliferation of the *Reference Test Arenas for Urban Search and Rescue Robots* as representative search and rescue environments, which allow objective performance evaluation of mobile robot capabilities. These arenas have hosted annual rescue robot competitions sponsored by the *American Association for Artificial Intelligence* since 2000 and the *RoboCupRescue Robot League* since 2001. The rules of these competitions have evolved each year to encourage robots to negotiate complex and collapsed structures, find simulated victims, determine their condition and location, and generate human readable maps to enable victim recovery. The associated performance metric has also evolved as it attempts to quantify and encourage these and other robot capabilities pertinent to urban search and rescue applications. This paper presents the competition rules, performance metric, and generalized results of the 2003 competitions which included some inspiring robotic implementations. Performance data captured during these competitions is discussed along with recently available development tools which can quicken the pace of innovation in the field of search and rescue robotics.

KEYWORDS: *mobile robot, autonomous robot, urban search and rescue robot, sensory perception, planning, mapping, collaboration, operator interface, robot test arenas and performance metric*

1. INTRODUCTION

The *Reference Test Arenas for Urban Search and Rescue Robots*, developed by the National Institute of Standards and Technology (NIST), test mobile robot capabilities and human/robot interfaces in arenas which are representative of buildings in various stages of collapse. They have been used to host annual urban search and rescue (USAR) robot competitions sponsored by the *American Association for Artificial Intelligence* (AAAI) since 2000, and the *RoboCupRescue Robot League* since 2001 [1] [2] [3] [4]. The goal of these competitions is to increase awareness of the challenges involved in search and rescue applications, provide objective evaluation of robotic implementations, and promote collaboration between researchers. These competitions require robots to negotiate the arena's complex and collapsed structures, find simulated victims, and generate human readable maps to enable victim recovery. The rules encourage robots to demonstrate their capabilities in mobility, sensory perception, planning, mapping, and practical operator interfaces, while searching for simulated victims. The performance metric attempts to quantify these and other robot capabilities pertinent to USAR applications for the purposes of comparison between diverse robotic implementations and team strategies.

Each year, the AAAI and RoboCupRescue competitions follow the same rules. These competition rules and the associated performance metric evolve as necessary to encourage technical advances, discourage certain unhelpful team strategies, and increase the operational relevance of successful teams to real world disaster situations. These year-to-year refinements attempt to guide researchers toward the following league vision:

When disaster happens, minimize risk to search and rescue personnel, while increasing victim survival rates, by fielding teams of collaborative robots that can:

- *Negotiate compromised and collapsed structures*
- *Find victims and ascertain their conditions*
- *Produce practical maps of the environment*
- *Deliver sustenance and communications*
- *Embed sensors and communication networks*
- *Identify hazards*
- *Provide structural shoring*

…allowing human rescuers to quickly locate and extract victims. [5] [6]

The arenas, competition rules, and the performance metric used during the 2003 competitions are presented below, along with a discussion of robot performance, inspiring implementations, and recently available development tools.

2. ARENA PROLIFERATION

The *Reference Test Arenas for Urban Search and Rescue Robots*, named *yellow*, *orange*, and *red* to indicate their increasing levels of difficulty, form a continuum of challenges for robots and operators [7]. A maze of walls, doors, and

Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE **SEP 2003**	2. REPORT TYPE	3. DATES COVERED **00-00-2003 to 00-00-2003**
4. TITLE AND SUBTITLE **Evolution of a Performance Metric for Urban Search and Rescue Robots (2003)**		5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **National Institute of Standards and Technology (NIST),100 Bureau Drive,Gaithersburg,MD,20899**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
PerMIS?03, Performance Metrics for Intelligent Systems, 16-18 Sep 2003, Gaithersburg, MD

14. ABSTRACT
see report

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**10**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

elevated floors provide various tests for robot navigation and mapping capabilities. Variable flooring, overturned furniture, and problematic rubble provide obvious physical obstacles. Sensory obstacles, intended to confuse specific robot sensors and perception algorithms, provide additional challenges. Intuitive operator interfaces and robust sensory fusion algorithms are highly encouraged to reliably negotiate the arenas and locate victims.

In 2003, competitions were held using existing and newly constructed arenas. For example, the first RoboCupRescue Japan Open competition was held in arenas fabricated for last year's RoboCupRescue2002 competition in Fukuoka, Japan [8]. A new orange arena was fabricated at Carnegie Mellon University and used to host demonstrations at the first RoboCupRescue U.S. Open. It will be used to support year-round robotics research. Also, new Italian arenas were fabricated for the RoboCupRescue2003 competition in Padua, Italy (Figure 1). These arenas will reside year round at the Istituto Superiori Antincendi in Rome, a fire-rescue training facility, and will support European robotics research. They may even host an Italian Open event next year.

Figure 1: RoboCupRescue Robot League Arenas (yellow, orange, and red)

The 2003 AAAI competition continued to use the NIST transportable arenas to host the competition in Acapulco, Mexico (Figure 2).

Figure 2: AAAI Mobile Robot Rescue Competition Arenas (yellow, orange, and red)

Currently, there are four year-round arena facilities in three countries around the world (see Table 1). Constructed to host previous competitions, they now provide ongoing support for research in this domain by raising awareness of the challenges facing robots and allowing practice in representative environments. More arenas are being planned for this year. Preparations are underway to fabricate new arenas to host the main RoboCupRescue2004 competition in Lisbon, Portugal. New arenas may also be fabricated in Germany and the U.S. to host Open competitions in 2004.

PREVIOUS COMPETITIONS	YEAR-ROUND ARENAS
2000 AAAI Conference AUSTIN, TX, USA	**NIST** MARYLAND, USA (2000)
2001 IJCAI/AAAI Conference SEATTLE, WA, USA	**Museum of Emerging Science** TOKYO, JAPAN (2002)
2002 RoboCupRescue FUKUOKA, JAPAN	**Carnegie Mellon University** PITTSBURGH, PA, USA (2003)
2002 AAAI Conference EDMONTON, CANADA	**Istituto Superiore Antincendi** ROME, ITALY (2003)
2003 RoboCupRescue - U.S. Open PITTSBURGH, PA, USA	
2003 RoboCupRescue - Japan Open NIIGATA, JAPAN	
2003 RoboCupRescue PADUA, ITALY	
2003 IJCAI/AAAI Conference ACAPULCO, MEXICO	

Table 1: Summary of Competitions and Locations of Year-Round Arenas 2000-2003

3. THE SIMULATED VICTIMS

The objective for each robot entering the arenas, and the incentive to traverse every corner of each arena, was to find simulated victims. Each simulated victim was a clothed mannequin emitting body heat and other signs of life including motion (shifting, waving), sound (moaning, yelling, tapping), and/or carbon dioxide to simulate breathing (Figure 3). Particular combinations of these sensor signatures implied the victim's state: unconscious, semi-conscious, or aware.

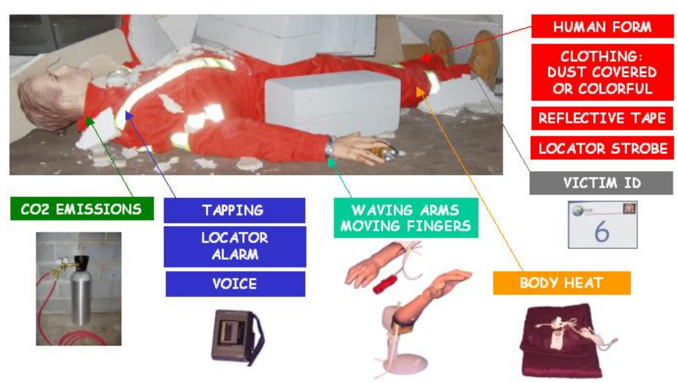

Figure 3: Simulated victims emit signs of life (surface victim shown)

Each victim was placed in a particular rescue situation, and distributed throughout the environment in roughly the same situational percentages found in actual earthquake

statistics: surface (50 %), lightly trapped (30 %), void (15 %), or entombed (5 %) [9] (Figure 4).

Figure 4: Other victim situations found in the arenas (trapped, void, and entombed)

The 2003 simulated victims remained similar to previous competitions, although a victim identification tag was introduced in 2003. These tags were usually placed in hard to reach areas around each victim and required advanced robot mobility to access and identify them. Tags were also placed in locations where victims were likely to be found, whether a victim was there or not, to inspire focused searches of the area with multiple sensors. This was meant to emulate a rescue dog handler's pointing motion, which helped to educate both robot researchers and the general public about targeted search techniques. Since penalties were assessed for false identifications, teams needed to be careful to identify appropriate signs of life along with any tag before reporting that a victim was found.

Generally, once a victim was found the robot(s) (currently with operator assistance) needed to determine the victim's location, situation, state, and identification tag, and then report their findings on a human readable map. These tasks formed the basis for the rules and performance metric discussed in the following sections.

4. The Competition Rules

These USAR competitions were classified as ranked competitions using objective scoring based on specified criteria [10] and the rules were developed by the joint AAAI/RoboCupRescue steering committee. The rules focused on the basic USAR tasks of setting up an operator station, safely negotiating the complex and collapsed structures in the arenas, allowing clear comparisons of diverse robotic implementations, and demonstrating reasonable operational stamina. Several refinements to the rules were introduced for the 2003 competitions to increase the operational relevance, simplify the judging, and add performance incentives [11].

The teams were given ten minutes to set up their operator station near the arenas prior to the mission start. The teams competed in missions lasting twenty minutes, with the winner achieving the highest cumulative score from seven (RoboCupRescue) or five (AAAI) missions depending on the event duration. The team's lowest mission score was dropped to allow one robot failure without consequence.

The first mission of each round began at the yellow arena entrance to allow direct comparison of capabilities across all teams. In later missions, teams were allowed to start their robots at the entrance to more advanced arenas. This allowed purpose-built robots to highlight their specific capabilities without retracing simpler arenas.

To encourage robot collaboration, teams with more than one robot were required to start at the same specified arena entrance and sequentially negotiate the arenas (yellow to orange to red). Any or all robots could advance as far as they wanted toward more difficult arenas. But multiple robot teams could not enter two arenas simultaneously from any start point, and needed to retreat as a team to enter simpler arenas.

If a robot became incapacitated or stuck, the operator could request a reset of the robot back to the start point. This allowed disabled robots to resume operation for the remainder of the mission, but penalized the team by adding one additional operator in the performance metric as discussed in the next section.

Several rule changes may be initiated for next year's competitions. For example, the ability to quickly set up an effective operator control station with a minimum of operators is a key requirement for deployment in actual disaster scenarios. This year, no penalties were enforced for taking longer than the allotted ten minutes to setup the operator station, and any number of team members could assist in this task. Only the people present in the operator station between mission start and end were counted as operators in the performance metric. Next year, both the time to set up and the number of team members required may be figured into overall team performance.

New rules or performance incentives may encourage more automated identification of a victim's signs of life. Operators typically identified potential victims from streaming video and then checked their additional sensors to verify other signs of life. Next year, more emphasis may be placed on the system's capacity to autonomously recognize signs of life and alert the operator to make visual verification. This may be accomplished by having signs of life in the arena with no obvious visual cues, or by modifying the rules and/or the performance metric to encourage such capabilities.

Also, to simulate the problematic radio interference and radio signal dropout that occurs at actual disaster sites, the rules may artificially interfere with radio communications during missions. The intent would be to encourage more development of autonomous behaviors, tether management systems, or other functional improvements that would benefit eventual deployment systems.

5. The Performance Metric

The performance metric used for scoring the competitions focused on the basic USAR tasks of identifying victims; determining their condition, situation and precise location; and enabling victim recovery through generation of human readable maps - all without hurting the victims or causing

$$\text{ARENA WEIGHTING}\left(\frac{\text{MAP QUALITY} + \text{VICTIM LOCATION} + \text{VICTIM TAG} + \text{VICTIM SITUATION} + \text{VICTIM STATE} - \text{ARENA BUMPING} - \text{VICTIM BUMPING}}{\left(1 + \text{NUMBER OF OPERATORS}\right)^2}\right)$$

Figure 5: The performance metric used for scoring competitions

damage to the environment (Figure 5). It particularly encouraged perception of detailed victim information through multiple sensors. Teams were also encouraged to minimize the number of operators, which could be achieved through use of better operator interfaces and/or autonomous behaviors that allowed effective robot control of multiple robots. Finally, arena weightings accounted for the difference in difficulty of negotiating each arena. The more difficult the arena, the higher the arena weighting (score) for each victim found.

Up to (50) points were available for each victim found based upon a variety of factors. However, points were also deducted for errant victim identifications or uncontrolled bumping of victims or arena features. The penalties were meant to encourage confidence in reported results and promote safe operation within dangerous environments. The performance metric's point allocations were as follows:

5.1. Map Generation

Up to (20) points per victim were available for generation of a paper-based map of the environment submitted by the end of each mission. The map was graded on a three-step scale (1, 5, 10 points) according to the following criteria:

- (10) points per victim were available for MAP QUALITY. The map needed to clearly identify found victims, discernable arena features and/or hazards, and any other helpful information to quickly deploy rescuers and appropriate tools required to extricate the victim. An automatically generated map with operator annotations received the maximum (10) points. Computer-generated maps that were operator interpreted or corrected received (5) points. Human-generated maps or topological maps received (1) point due to the heavy workload placed on the operator.

- (10) points per victim were available for VICTIM LOCATION. The map needed to accurately locate found victims. Any part of a found victim identified to within 1 cubic meter of a recognizable reference point or arena feature received the maximum (10) points. Locating the victim to an adjacent cubic meter (not through walls) received (5) points. Locating the victim to any other location received (1) point.

5.2. Sensory Perception

Up to (15) points per victim were available for identification of the VICTIM STATE through individual sensory perception capabilities and correct interpretation of particular sensor combinations (Figure 3).

- (10) points per victim were available for interpreting various sensor signatures as potential signs of life. The operators were allowed to visually interpret video images to determine human form and motion, and otherwise interpret sensor signatures to determine heat, sound, and CO_2 (average amount per breath exhaled by an adult). Automatic identification of valid sensor signatures would be clearly desirable and will likely be further encouraged in future iterations. Misinterpreting sensor signatures, or false positive identifications, resulted in negative points. The points available for specific sensor signatures were as follows:

HUMAN FORM (head/torso, arm, legs, baby)	(+/- 1 point)
MOTION (none, shifting, waving,)	(+/- 1 point)
BODY HEAT (37 °C ± 2 °)	(+/- 3 point)
SOUND (none, moaning, yelling, tapping)	(+/- 2 point)
CO_2 (35,000ppm to 50,000ppm)	(+/- 3 point)

- (5) points per victim were available for correctly discerning the VICTIM STATE from at least three different sensor signatures. Particular combinations of sensor signatures described a victim state as unconscious, semi-conscious, or aware. For example, an operator who correctly identified body heat, arm-waving, and audible yelling as an "aware" victim received (5) points. Incorrectly identified states received no penalty since the concept was new to the competition (this may change in the future).

5.3. Mobility

Up to (15) points per victim were available for demonstration of advanced mobility via remote teleoperation or autonomous control modes.

- (5) points per victim were available for remotely discerning the VICTIM SITUATION by being mobile enough to circumnavigate the victim or access an advantageous viewing position. The identifiable situational categories for victims were surface, lightly trapped, void, or entombed. Generally, surface victims were entirely visible. Lightly trapped victims were partially visible, requiring assistance to remove rubble pinning them in position. Victims trapped in voids were slightly visible under leaning or pancake collapses requiring much more effort to extricate. Entombed victims were not directly visible but emitted signs of life from under large boxes or massive rubble (Figure 4).

- (10) points per victim were available for remotely reading the VICTIM TAG located on or near the victim. These tags were placed in hard-to-reach areas around the victim and required advanced robot mobility and reasonable quality image resolution to identify. Accessibility to awkward or confined spaces and good quality images are essential to gain confidence in a finding and to communicate pertinent details to medical personnel or other experts. A misread tag resulted in (–10) points, so close proximity to the tag and high confidence in the reported information was encouraged.

5.4. Penalties

As discussed above, the performance metric contains point reductions to discourage false victim identifications. In addition, penalties for uncontrolled bumping behaviors were imposed in two levels of severity. Light bumping of victims or arena features received (-5) points per incident. Hurting simulated victims or causing secondary collapses received (-20) points per incident. Penalties could compound.

5.5. Operators

Any person who entered the operator station during a mission was considered an operator. This encouraged a minimum of operators required to perform the given tasks. The intent also encouraged an increase in the ratio of robots to operators either through demonstration of effective autonomy (sliding, bounded, or total), or intuitive operator interfaces that allowed high-level management of multiple robots.

5.6. Arena Weighting

Arena weighting factors accounted for the difference in difficulty negotiating each arena; the more difficult the arena, the higher the arena weighting (score) for each victim found. The intent was to leverage the increased time required to negotiate more difficult terrain and perceive the simulated victims in more complex environments. The arena weighting factors were (1.0) for the red arena, (0.75) for the orange arena, and (0.5) for the yellow arena. So finding red arena victims counted for twice that of finding yellow arena victims and encouraged teams to attempt the more difficult arenas.

6. COMPETITION RESULTS

The RoboCup2003 Rescue Robot League competition hosted twelve teams that demonstrated robotic systems with very diverse characteristics (Figure 6). The first place award winner was the ROBRNO team from Brno University of Technology in the Czech Republic [12]. They developed a very capable custom robot and integrated several components to form an extremely effective operator interface. Their robustly fabricated four-wheel, skid-steered robot was equipped with vision, infrared, and audio sensors for victim identification. The operator interface included a joystick to control robot motion along with heads-up display goggles that tracked the orientation of the operator's head to automatically point the robot's cameras. This allowed superior remote situational awareness and enabled the operator to negotiate narrow arena passages intuitively and dexterously, causing very few penalties.

Figure 6: Robots from RoboCup2003 Rescue

The second place award winner was the CEDRA team from Sharif University of Technology in Iran [13]. They developed a wheeled mobility platform with an articulated body design similar to planetary explorers. They also employed a joystick interface with the operator viewing two flat panel video displays. The third place award winner was the MICROBOT team from the Isfahan University of Technology (IUT) also in Iran [14]. They demonstrated two robots equipped differently and used cooperatively. One robot was small and fast with only a camera for initial victim identification and operator-generated mapping. Once a victim was potentially located, the second, slower robot was dispatched to the location with more specific victim identification sensors. The technical award winner was the team from the International University - Bremen (IUB) in

Germany [15]. They also deployed two robots, but were recognized for their mapping implementation, which used a proximity range finder to automatically generate obstacle maps of the environment. This was the only autonomous mapping demonstrated during the competition which was highly encouraged in the performance metric, yet did not contribute quite enough points for them to earn a place award.

Other interesting approaches included fully autonomous robots, a robot almost directly from the mid-size soccer league, and even a blimp. The two fully autonomous teams demonstrated robots capable of navigating parts of the yellow arena but did not produce maps showing victim identifications, another key performance criteria, so these systems did not score well. The remotely teleoperated teams showed few autonomous behaviors to assist their operator's efforts, although several teams were working toward such capabilities. Most teams used wireless communications between the robots and their operator station, while a few teams used fixed tethers with varying levels of success due to snagging obstacles in the environment.

Figure 7: Robots from IJCAI/AAAI2003 Rescue

The IJCAI/AAAI-2003 Rescue Robot competition hosted six teams demonstrating both commercially available robots with enhanced control and low-cost autonomous robots (Figure 7). The first place team was the Idaho National Engineering and Environmental Laboratory (INEEL) team from the USA. This team used a commercially available ATRV-Jr. equipped with vision, sonar, infrared, and laser sensors to explore the yellow and orange arenas. They featured a comprehensive operator interface that displayed sensor readings, robot status information and environmental maps. But they were most noted for their approach toward sliding autonomy that allowed the operator to choose varying degrees of control from pure teleoperation to full autonomy at any point during a mission. This capability clearly reduced the operator's workload and greatly assisted in negotiating narrow passages.

The second place team was Swarthmore College from the USA. This team used a single operator to deploy two robots with varying degrees of autonomy. Their mapping implementation allowed the operator to tag interesting points in a robot view of the environment and use them to compensate for robot position errors. They also demonstrated a web-based victim information form that the operator used to capture and convey pertinent location, situation, and state information for each victim.

The technical award winner was the University of New Orleans (UNO) from the USA. They deployed four Sony Aibo dogs and a blimp, but were recognized for their collaborative mapping approach toward building a graphical 3-D model of the environment noting walls, obstacles and victim locations.

Other interesting systems included two low-cost but fully autonomous robots. Both teams focused on the low end of the cost spectrum in an attempt to field swarms of similar robots to explore unstructured environments. They explored various parts of the yellow arena, but were unable to identify victims or produce maps, two key factors in the performance metric, so they did not score well. Another team teleoperated a robot with a fixed tether, but had limited success exploring the yellow and orange arenas.

7. PERFORMANCE DATA CAPTURED

7.1 Robot Video and Operator interfaces

For the second year, human-factors researchers from NIST, the University of Massachusetts, and the Mitre Corporation used the competition event to study human-robot interaction during missions [16] [17] [18]. The operators, the interfaces to their robots, and the robots themselves were all video taped during missions. These video streams, objective monitoring of operator actions, and interviews conducted immediately after each mission captured the workload required to perform each task and provided the basis for study of operational effectiveness and ease-of-use issues. A formal analysis of this data is underway with the goal of identifying particularly effective interface components and methods so that other teams and applications may benefit.

7.2 Performance Statistics

Also at this year's competitions, several pertinent performance statistics were captured and tabulated for analysis (Figure 8). This performance data allowed comparisons among teams within the same competition, across competitions within the same year, and provided a performance baseline to gauge improvements year to year. The goal is to use statistically significant performance data to quantify the state-of-the-art in USAR robots while highlighting clear performance improvements. Some notable statistics provided a few insights. For example, the average scores, high scores, average number of penalties, and average points per penalty of this

year's RoboCup and AAAI competitions were comparable, which suggested that both sets of arenas posed similar levels of difficulty as intended. Also, the top teams in each competition scored well not necessarily because they found a greater average number of victims, but because they scored a higher average of points per victim. The AAAI teams used fewer operators per team on average than the RoboCupRescue teams, which suggested either more autonomy or better operator interfaces. Although, the AAAI teams also required

2003 Statistics	No. of Missions	High Score	Average Score	Operators (Avg.)	Resets (Avg.)	Avg. No. of Victims Found (Avg. Points per Victim)			Avg. No. of Penalties (Avg. Points per Penalty)			Missions > 0 (%)
						Yellow	Orange	Red	Yellow	Orange	Red	
RoboCupRescue (Padua, Italy, July 2003)												
Team A	7	23.8	15.1	1.0	0.0	1.1 (11.5)	3.0 (18.0)	1.3 (14.0)	0.4 (10.0)	0.3 (5.0)	0.3 (5.0)	100%
Team B	7	12.5	7.6	1.0	0.0	1.3 (10.8)	2.1 (8.8)	1.1 (9.6)	0.4 (5.0)	0.0 (0.0)	0.1 (5.0)	100%
Team C	7	7.3	3.0	1.0	0.0	2.1 (9.3)	1.1 (9.5)	0.0 (0.0)	1.4 (5.0)	0.9 (5.0)	0.0 (0.0)	86%
Team D	7	5.9	2.4	1.0	0.1	1 (17.0)	0.3 (12.0)	0.0 (0.0)	0.1 (5.0)	0.0 (0.0)	0.0 (0.0)	71%
Team E	5	13.2	5.0	1.6	0.6	3.6 (13.1)	2.7 (9.6)	0.0 (0.0)	0.8 (5.0)	1.2 (5.0)	0.0 (0.0)	80%
Team F	5	5.6	2.1	1.0	0.4	0.6 (15.0)	1.7 (12.2)	0.0 (0.0)	0.0 (0.0)	0.2 (5.0)	0.0 (0.0)	40%
Team G	3	1.8	0.7	1.3	0.7	1.7 (6.8)	0.0 (0.0)	0.0 (0.0)	1.0 (5.0)	0.0 (0.0)	0.0 (0.0)	67%
Team H	3	1.5	0.5	1.0	0.7	0.7 (8.5)	0.0 (0.0)	0.0 (0.0)	0.7 (5.0)	0.0 (0.0)	0.0 (0.0)	67%
Team I	3	0.8	0.3	1.0	0.0	0.7 (12.3)	0.0 (0.0)	0.0 (0.0)	1.3 (5.0)	0.0 (0.0)	0.0 (0.0)	67%
Team J	3	0.0	0.0	1.0	0.3	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	1.0 (5.0)	0.0 (0.0)	0.0 (0.0)	0%
Team K	3	0.0	0.0	2.0	0.0	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0%
Team L	3	Technical Difficulties - Did not compete										-----------
SITE AVGS.	**4.8**	**6.6**	**3.3**	**1.2**	**0.3**	**1.2 (9.5)**	**1 (6.4)**	**0.2 (2.2)**	**0.6 (4.5)**	**0.2 (1.8)**	**0.0 (0.9)**	**61.6%**
IJCAI/AAAI (Acapulco, Mexico, August 2003)												
Team M	5	27.6	9.2	1.0	0.0	1.8 (24.3)	1.6 (21.7)	0.0 (0.0)	1.0 (8.0)	2.6 (6.2)	0.0 (0.0)	100%
Team N	5	12.5	6.1	1.0	0.2	2.4 (22.9)	0.0 (0.0)	0.0 (0.0)	0.4 (12.5)	0.0 (0.0)	0.0 (0.0)	60%
Team O	5	2.9	1.6	2.4	0.0	2.2 (12.6)	0.4 (10.5)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	100%
Team P	5	4.0	1.5	1.2	0.2	1.8 (11.9)	0.0 (0.0)	0.0 (0.0)	2.0 (9.5)	0.6 (10.0)	0.0 (0.0)	40%
Team Q	3	0.0	0.0	0.0	2.0	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0%
Team R	3	0.0	0.0	0.0	0.3	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0%
SITE AVGS.	**4.3**	**7.8**	**3.1**	**0.9**	**0.5**	**1.4 (11.9)**	**0.3 (5.4)**	**0.0 (0.0)**	**0.6 (5.0)**	**0.5 (2.7)**	**0.0 (0.0)**	**50.0%**

Figure 8: Statistics captured during 2003 competitions

more resets to free stuck robots in order to perform the tasks. Another key statistic was the general productivity and reliability of each team, tracked as a percentage of missions producing a positive score ("Missions > 0"). Although not accounted for as part of the performance metric, reliability is certainly a key component of overall effectiveness.

7.2 Position Tracking of Robots

To augment these data collection efforts in future events, a robot tracking system is being developed at NIST to automatically capture robot position during missions. Objective position tracking will provide a powerful tool for comparing the effectiveness of diverse robotic implementations by capturing statistics like arena coverage, search rates, dwell times, proximity to victims, and other key measures. Tracking will allow further statistical analysis of robot performance to clearly identify successful implementations while revealing problem areas as well. When combined with the video capture of robot performance and operator workload, this tracking system will provide valuable feedback to robot researchers and project sponsors. Initially, this tracking system will be used to capture performance data during robot practice sessions in the NIST arenas and at the 2004 Rescue Robot League competitions. Eventually, the robot tracking system may be replicated for other test arenas around the world.

8. ASSOCIATED DEVELOPMENT TOOLS

Virtual versions of NIST's *Reference Test Arenas for Urban Search and Rescue Robots* have been developed to provide the research community with an efficient way to test their algorithms without having to incur the costs associated with maintaining functional robots and traveling to one of the permanent arena sites for validation and practice. These virtual arenas consist of real sensor data sets and simulated environments.

8.1 Sensor Data Sets

The sensor data sets allow programmers anywhere to access and use a variety of actual sensor data readings gathered inside the NIST arenas. A systematic (1m x 1m) grid under the yellow arena allows precise placement of sensors over grid nodes to represent robot sensor views facing north, east, south, and west inside the arena. These sensor readings are compiled

into a database allowing easy extraction for perception, planning, or other sensor-based algorithms.

So far two sensors have been used, a digital camera and a line-scan LADAR, but other advanced (even developmental) sensors are being considered such as the so-called “flash LADAR.” In this way, a single sensor can produce situational data sets for anybody interested in algorithm testing, even before the sensors are widely available or cost effective. The resulting sensor-based algorithms, when shown to be effective in navigating the virtual data sets, should have a high likelihood of success when the actual sensor is deployed in the real arenas. This may occur during competitions or other testing opportunities in a permanent arena.

These databases are available in both Linux and Windows formats. They allow the user to export captured data from a GDBM database to useable files that may be inserted into algorithms for testing and development. A Matlab graphical interface supplements the query databases and allows the user to manually select and view data from any grid node within the arena. This is useful to peruse the data in search of particular data samples, or to troubleshoot failures at specific locations.

8.2 Arena Simulations

The University of Pittsburgh and Carnegie Mellon University recently developed a realistic simulation of the *Reference Test Arenas for Urban Search and Rescue Robots* using a game engine graphics environment [19]. This pseudo-dynamic simulation of the NIST orange arena supports hardware-independent algorithm development with simulated sensor signatures, and adds the ability to virtually design and test new robotic mechanisms and sensor configurations. The popularity of the underlying game engine, Unreal Tournament, will hopefully capture the imagination of programmers and may entice more people into developing robots and capabilities applicable to search and rescue. Currently, there is one robot modeled for use in the environment, but the tools are available to design and test other ideas. Although only range sensing is currently available, work is continuing toward simulating a line-scan LADAR and infra-red type heat sensor. Progress is also being made to add the other arenas to the simulation.

By minimizing the jump from the virtual to the real world arenas, and allowing hardware-independent testing of concepts, this simulation tool will hopefully quicken the developmental pace of more capable systems.

8.2 Reality Arena

Originally an underground missile silo, and more recently a stairwell burn-facility, the new “reality” arena (also known as the black arena) has been converted into a hardened, difficult, robot test facility that is safe for researchers and robots alike (see Figure 9).

Figure 9: Reality Arena (above, below ground images)

This NIST facility contains several above ground features (doors, windows, fire-escape, and skylight entrance), but the majority of the 300 square-meter facility is located underground as a reinforced concrete and steel structure. Inside this facility, the same simulated victims are placed in stairwells, narrow passages, confined spaces, and under collapsed debris (see Figure 10). This arena is dirty, difficult, even wet at times, adding a sense of realism for the most competent robots. Operators find this arena the most challenging by far because it is too vast to understand intuitively. Lacking a human level understanding of the environment provides the clearest indication of where operator interfaces fail to provide sufficient situational awareness. Meanwhile, radio communications in this environment are also problematic, just as at actual structural collapses. So this reality arena provides a comprehensive test for the most capable systems and is available year-round to researchers.

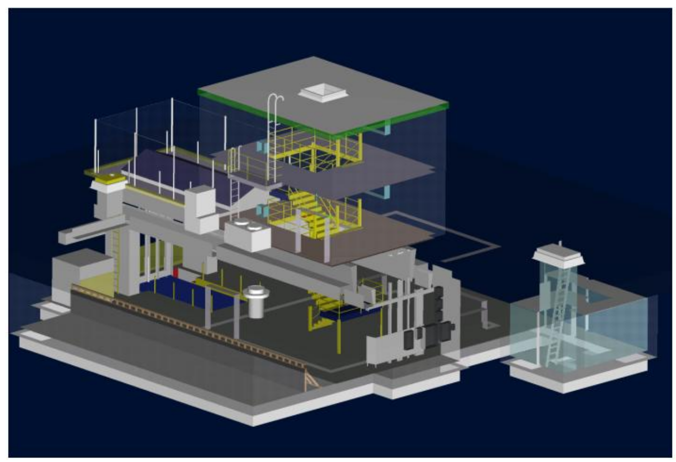

Figure 10: Reality Arena (transparent 3D-CAD model)

9. FUTURE DIRECTIONS

The *Reference Test Arenas for Urban Search and Rescue Robots*, and the annual USAR robot competitions, allow direct comparison of robotic approaches and encourage objective performance evaluation. As robot teams begin demonstrating repeated successes against the obstacles posed in the arenas, the level of difficulty will be increased accordingly so that the arenas provide a necessary stepping-stone from the laboratory

toward eventual robot deployment in real disaster sites. New arenas are being constructed this year to support the upcoming competitions. Afterward, they will provide year-round practice facilities in those countries to support research and development. The upcoming competitions are:

RoboCupRescue - U.S. Open
NEW ORLEANS, LA, USA (APRIL 24-27, 2004)

RoboCupRescue - Japan Open
OSAKA, JAPAN (MAY 1-4, 2004)

RoboCupRescue - German Open
PADERBORN, GERMANY (TBD)

RoboCupRescue
LISBON, PORTUGAL (June 29 - July 3, 2004)

AAAI Conference
SAN JOSE, CA, USA (July 25 - 29, 2004)

The performance metric used for scoring these competitions will also evolve as necessary to encourage application of pertinent technologies according to the league vision. New virtual arenas were introduced that provide both sensor data sets and simulated dynamic environments. They will allow hardware-independent development of algorithms and testing for innovative robot designs. These virtual development tools may eventually provide a point of collaboration between the RoboCupRescue Simulation League and the RoboCupRescue Robot League. Meanwhile, efforts to capture robot, operator, and team performance during year-round testing sessions at NIST and at the yearly competitions will continue so that progress may be documented and advances more readily adopted. The addition of an automatic robot tracking system will provide valuable performance feedback to robot researchers. As these data collection efforts document the state of the art in robotic capabilities, the yearly competitions will continue to provide public proving grounds for field-able robotic systems that will ultimately be used to save lives.

10. NIST Disclaimer

Commercial equipment and materials are identified in order to adequately specify certain procedures. In no case does such identification imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the material or equipment identified are necessarily the best available for the purpose.

11. Acknowledgements

This work was funded in part by the DARPA Mobile Autonomous Robot (MARS) Program.

12. References

[1] Schultz, A. "The 2000 AAAI Mobile Robot Competition and Exhibition," *AI Magazine*, 22(1), spring 2001, AAAI.

[2] Murphy, R., Blitch, J., and Casper, J. "AAAI/RoboCup-2001 Urban Search and Rescue Events," *AI Magazine*, 23(1): Spring 2002, 37-42.

[3] Casper, J. and Yanco, H. "AAAI/RoboCup-2001 Robot Rescue," *AI Magazine*, 23(1): Spring 2002, 43-50.

[4] Asada, M. et al. "An Overview of RoboCup-2002 Fukuoka/Busan" *AI Magazine*, 24(2): Summer 2003.

[5] Jacoff, A., Messina, E., Weiss, B., Tadokoro, S., and Nakagawa, Y. "Test Arenas and Performance Metrics for Urban Search and Rescue Robots," *International Conference on Intelligent Robots and Systems*," October 2003, Las Vegas, NV.

[6] Blitch, J. "Robot-Assisted Marsupial Deployment of Urban Unattended Ground Sensors," *Proceedings of SPIE Unattended Ground Sensor Technologies and Applications V Conference*, Vol. 5090, Orlando, Florida, April, 2003.

[7] Jacoff, A., Messina, E., and Evans, J. "A Standard Test Course for Urban Search and Rescue Robots," *Performance Metrics for Intelligent Systems Workshop*, August 2000, Gaithersburg, MD.

[8] Jacoff, A., Messina, E., and Weiss, B. "Intelligent systems for urban search and rescue: challenges and lessons learned," *SPIE's AeroSense-Aerospace/Defense Sensing, Simulation, and Controls Conference*, April 2003, Orlando, FL.

[9] Murphy, R., Casper, J., Micire, M. and Hyams, J. "Assessment of the NIST Standard Test Bed for Urban Search and Rescue," *Performance Metrics for Intelligence Systems Workshop*, August 2000, Gaithersburg, MD.

[10] Yanco, H. "Designing Metrics for Comparing the Performance of Robotic Systems in Robot Competitions," *Performance Metrics for Intelligent Systems Workshop*, September 2001, Mexico City, Mexico.

[11] "RoboCup and AAAI Rescue Robot Competition Rules," *http://robotarenas.nist.gov/competitions.htm*, 2004.

[12] Zalud, L. "RoboCup Rescue Robot League Competition Awardee Paper – Robrno," *RoboCup2003: Robot Soccer World Cup VII – Lecture Notes on Artificial Intelligence,*

edited by: Polani D., Browning, B., Bonarini, A., and Yoshida, K., Springer, 2004.

[13] Meghdari, A., Amiri, F., Baghani, A., Mahboubi, H., Lotfi, A., Khalighi, Y., Karimi, R., Nejat, H., Amirian, M., Kamali, Sh., and Moradi, S. "RoboCup Rescue Robot League Competition Awardee Paper – Cedra," *RoboCup2003: Robot Soccer World Cup VII – Lecture Notes on Artificial Intelligence*, edited by: Polani D., Browning, B., Bonarini, A., and Yoshida, K., Springer, 2004.

[14] Mirmohammad Sadeghi, H., Bastani, H., and Azarnasab, E. "RoboCup Rescue Robot League Competition Awardee Paper – IUTMicrobot," *RoboCup2003: Robot Soccer World Cup VII – Lecture Notes on Artificial Intelligence*, edited by: Polani D., Browning, B., Bonarini, A., and Yoshida, K., Springer, 2004.

[15] Carpin, S., Kenn, H., and Birk, A. "Autonomous Mapping in Real Robots Rescue League," *RoboCup2003: Robot Soccer World Cup VII – Lecture Notes on Artificial Intelligence*, edited by: Polani D., Browning, B., Bonarini, A., and Yoshida, K., Springer, 2004.

[16] Drury, J., Scholtz, J., and Yanco, H. "Awareness in Human-Robot Interactions" *IEEE*, 2003.

[17] Yanco, H., Drury, J., and Scholtz, J. "Beyond Usability Evaluation: Analysis of Human-Robot Interaction at a Major Robotics Competition" *To appear in the Journal of Human-Computer Interaction*, 2004.

[18] Scholtz, J., Young, J., Drury, J., and Yanco, H. "Evaluation of Human-Robot Interaction Awareness in Search and Rescue," *Submitted to ICRA 2004.*

[19] Wang, J., Lewis, M., and Gennari, J. "A Game Engine Based Simulation of the NIST Urban Search and Rescue Arenas." *Proceedings of the 2003 Winter Simulation Conference*